

Drop Preference Services

Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

To view the entire list of current Internet-Drafts, please check the "1id-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), ftp.nordu.net (Northern Europe), ftp.nis.garr.it (Southern Europe), munnari.oz.au (Pacific Rim), ftp.ietf.org (US East Coast), or ftp.isi.edu (US West Coast).

Abstract

Drop preference has been proposed as a possible means for providing differentiated services in the Internet. Used as a parameter in a PHB group, drop preference can provide weighted bandwidth sharing amongst flows and flow aggregates of a PHB group. This memo documents a number of services which can be implemented using either a single drop preference PHB setting or multiple drop preference settings.

1. Introduction:

Drop preference (DP) has been proposed in both connection-oriented [[ATM94](#),[ATM96](#)] and connectionless networks [[Diffserv](#)] as a means to provide service differentiation between individual flows and flow aggregates. The idea behind drop preference is to mark packets with different drop priorities and to selectively drop lower priority packets when network resources are congested. In all of its proposed forms, drop preference is used to provide some form of weighted bandwidth sharing between flows in the network. Depending on how packets are marked, a variety of services can be implemented with this functionality. This memo documents a number of proposed services which can be implemented using either a single or multiple

drop preference settings/parameters within a PHB group. Services implementable using the expedited forwarding (EF) PHB [[Nichols](#)] are beyond the scope of this document.

2. Differential Services Architecture

Drop preference marking, in its current diffserv context, is done at the network edges using traffic conditioners (markers) which set the appropriate PHB fields. In the differential services framework, drop preference can be used as a parameter within a PHB group in order to provide service differentiation between flows of a particular PHB group. Within the network, the PHB setting specifies a local forwarding treatment or mechanism for transmitting the packet. While the proposed mechanisms and algorithms for providing drop preference functionality differ, most of them share the same basic characteristics [[Clark95](#), [Clark97](#), [Feng97](#), [Feng97a](#), [Kilikki](#)]. In particular, the mechanisms rely on queueing decisions at the bottleneck link in times of congestion. When such situations occur and queues build up, packets with lower priorities are dropped preferentially in an attempt to reduce the offered load at the bottleneck. In most cases, it is assumed that the packets of individual flows, regardless of their DP setting, are not re-ordered by the router. The use of drop preference gives the network a simple, low-overhead, fine-grained means to control bandwidth usage. By controlling the amount of priority marking done, the network can effectively control the bandwidth usage over time for individual flows and flow aggregates.

3. Two-level Drop Preference Services

There have been several proposed services which use a two-level drop preference setting as a building block in their implementations.

3.1 Assured and Controlled-load Services

In both assured and controlled-load service, the marker/policer determines the drop preference setting based on a fixed profile. In assured service [[Clark95](#), [Clark97](#)] the profile is based on the expected capacity of a flow/aggregate in times of congestion. The profile itself uses a fairly long sliding window to measure bandwidth usage. Based on this profile, packets within their expected capacity profile are marked as "In" (conformant) and given a higher priority. When congestion occurs and queues build up, low priority packets are dropped first while high priority packets are still allowed in the queue. In controlled-load service [[Wroclawski](#)], and in particular its diffserv implementation [[Feng97](#)], the profile is based on a token bucket specification of the flow or flow aggregate. Packets which are within the traffic envelope, as specified by the token bucket, are sent with higher priority while excess traffic is sent along with normal traffic with lower priority. In both cases, because of the burstiness of the traffic being policed and the dynamics of TCP congestion control, a large profile must be used in order to maintain predictable quality

of service.

Feng

Expires 12/98

[Page 2]

The marking behavior of assured and controlled-load service affords flows a constant amount of marking regardless of their sending rate at all times. That is, when a flow/aggregate sends an arbitrary amount over its profile, the marker still marks packets at the rate indicated in the profile. This marking differentiates the two services from SIMA and CBR as described below.

[3.2](#) Two-level SIMA (Simple Integrated Media Access)

The original SIMA service [[Kilkki](#)], as described in [Section 4](#), uses multiple drop preference levels to provide service differentiation between flows in times of congestion. In SIMA, sources contract a sending rate with the network referred to as a nominal bit rate (NBR). This rate serves as a profile from which drop preference marking is done. Depending on the actual rate the source sends, the network marks each packet of a flow with a particular drop preference setting which is calculated as a function of the profile and the actual sending rate. When the source sends at rates below its profile, its packets are marked with increasingly higher priority. In contrast, when the source sends at rates above its profile, its packets are marked with increasingly lower priority. The effect of this type of SIMA marking is to ensure that flows within their contracted rate receive a low amount of loss when congestion occurs. Flows which increasingly send above their contracted rate see increasingly higher packet loss rates in times of congestion as the SIMA marker changes lowers their priorities. While SIMA relies on using multiple drop priorities, it is possible to build a SIMA-like service using a two-level drop preference scheme. Instead of marking all of a flow's packets with a particular drop priority, a variable fraction of a flow's packets can be marked with either high or low priority. Flows sending below the profile have most of their packets marked with high priority. As flows send increasingly above their profile, the amount of high priority marking given steadily decreases to zero.

The marking behavior of this implementation of SIMA is different than both assured service and controlled-load service in that the amount of marking steadily decreases as the flow/aggregate exceeds its profile. As described above, the marking remains constant in assured and controlled-load service regardless of the amount of data the sources are sending above the profile.

3.3 CBR

Drop preference can also be used to implement a constant bit rate service [[Feng97a](#)]. As in assured service, the marker/policer observes the bandwidth usage of a connection using a sliding window and uses the contracted bit rate as a profile for determining its marking behavior. The marker implementing CBR service attempts to mark a minimal amount of packets in order for the flow/aggregate to meet its contracted profile. Thus, for a given profile, the CBR marker marks packets at a rate which is less than or equal to that of the assured and controlled-load services. The marker does this by observing the sending rate of the flow/aggregate over the sliding window. As long as the flow/aggregate is sending above its profile, the marker slowly reduces its marking. As long as the flow/aggregate is sending below its profile, the marker slowly increases its marking. Over a period of time, such adaptive marking delivers the flow/aggregate a constant amount bandwidth. By minimizing the amount of priority packets in the network, the CBR marker can help improve the performance of priority-based queueing algorithms in the network.

The marking behavior of CBR is different than the above services in that marking is reduced to zero when the flow/aggregate exceeds its profile. This is in contrast to the assured and controlled-load services which maintain a constant rate of marking regardless of the flow/aggregate's actual sending rate and in contrast to SIMA which slowly reduces the marking as the flow/aggregate increases its rate.

4. Multi-level Drop Preference Services

Several services which use more than two drop preference levels have been proposed to provide weighted bandwidth sharing amongst flows and flow aggregates.

4.1 SIMA

As described earlier, the original SIMA service uses multiple drop priorities to provide service differentiation between flows. In this service, marking is done strictly based on the sending rate of the flow/aggregate and its negotiated rate (NBR). The SIMA marker marks all of the packets of a particular flow/aggregate with the same priority according to how much it is sending with regard to its profile. As the source(s) increasingly sends above its profile, the marker steadily decreases the drop priority.

Feng

Expires 12/98

[Page 4]

4.2 Proportional Sharing

Cooperative Dropping [[Weiss](#)] is a framework for flexibly providing a number of bandwidth services (including proportional sharing) using multiple drop priorities. In this scheme, the packets of a flow/aggregate are striped across the multiple priority levels. At each intermediate router, preferential dropping is done depending on the priority level of the packets and the current level of congestion. Given an incoming multi-priority traffic stream, the congested router adaptively selects a drop priority threshold in which all packets below the threshold are dropped. Depending on how the packets of a flow/aggregate are striped across the multiple priorities, proportional sharing and a variety of other services can potentially be implemented. For example, using the preferential dropping mechanism in conjunction with a marker which stripes a flow/aggregate's packets evenly across all priorities, sharing in direct proportion to the flow/aggregate's sending rate can be achieved. Note that the granularity and accuracy of the bandwidth sharing is dependent on the number of priority levels used. Using more priority levels results in a finer degree of bandwidth sharing.

The marking behavior of proportional sharing, given a profile per priority level, remains fixed. Additional marking policies in the cooperative dropping framework can be deployed, however, in order to implement a larger range of services.

5. References

- [ATM94] ATM User-Network Interface (UNI) Signalling Specification Version 3.1 (af-uni-0010.002), September 1994.
- [ATM96] ATM Traffic Management Specification Version 4.0 (af-tm-0056.000), April 1996.
- [Clark95] D. Clark, "Adding Service Discrimination to the Internet",
<http://ana-www.lcs.mit.edu/anaweb/ps-papers/TPRC2-0.ps>,
September 1995.
- [Clark97] D. Clark and J. Wroclawski, "An Approach to Service Allocation in the Internet", Internet Draft
<[draft-clark-diff-svc-alloc-00.txt](#)>, July 1997.
- [Feng97] W. Feng, D. Kandlur, D. Saha, and K. Shin,
"Understanding TCP Dynamics in an Integrated Services Internet", Proc. of NOSSDAV '97,
<http://www.eecs.umich.edu/~wuchang/ered/> ,
July 1997.

- [Feng97a] W. Feng, D. Kandlur, D. Saha, and K. Shin, "Adaptive Packet Marking for Providing Differentiated Services in the Internet", UM CSE-TR-347-97, <http://www.eecs.umich.edu/~wuchang/pmg/> , September 1997.
- [Ferguson] P. Ferguson, "Simple Differential Services: IP TOS and Precedence, Delay Indication, and Drop Preference, Internet Draft <[draft-ferguson-delay-drop-00.txt](#)>, November 1997.
- [Diffserv] Differentiated Services Framework Draft, May 1998.
- [Kilikki] K. Kilikki, "Simple Integrated Media Access (SIMA)", Internet Draft <[draft-kalevi-simple-media-access-01.txt](#)>, June 1997.
- [Nichols] K. Nichols and S. Blake, "Definition of the Differentiated Services Field (DS Byte) in the IPv4 and IPv6 Headers", Internet Draft <[draft-ietf-diffserv-headers-00.txt](#)>, May 1998.
- [Weiss] W. Weiss, "Providing Differentiated Services through Cooperative Dropping and Delay Indication", Internet Draft <[draft-weiss-cooperative-drop-00.txt](#)>, March 1998.
- [Wroclawski] J. Wroclawski, "Specification of the Controlled-Load Network Element Service." [RFC 2211](#), September 1997.

Author's Address

Wu-chang Feng
University of Michigan
Dept. of EECS
Real-time Computing Laboratory
1301 Beal Ave./2228 EECS Building
Ann Arbor, MI 48109-2122
Phone: +1-734-763-5363
Fax: +1-734-763-4617
E-mail: wuchang@eecs.umich.edu

